Frequent Turbulence? A Dynamic Copula Approach

Lorán Chollete and Andreas Heinen *

July 31, 2006

Abstract

How common and how persistent are turbulent periods? We address these questions by developing and applying a dynamic dependence framework. In order to answer the first question we estimate an unconditional mixture model of normal copulas, based on both economic and econometric justification. In order to answer the second question, we develop and estimate a hidden markov model of copulas, which allows for dynamic clustering of correlations. These models permit one to infer the relative importance of turbulent and quiescent periods in international markets. Empirically, the three most striking findings are as follows. First, for the unconditional model, turbulent regimes are more common. Second, the conditional copula model dominates the unconditional model. Third, turbulent regimes tend to be more persistent.

Keywords: International Markets, Turbulence, Hidden Markov Model, Copula

*Chollete is at the Norwegian School of Economics and Business Administration (NHH), Heinen is at Universidad Carlos III, Madrid. Corresponding author is Chollete (loran.chollete@nhh.no). We are grateful for comments and suggestions from Jonas Andersson, Ricardo Caballero, Victor de la Peña, Andrew Gelman, Bruno Gerard, Philipp Hartmann, Andreas Heinen, Chris Heyde, Bob Hodrick, Jostein Lillestol, Ching-Chih Lu, Thomas Mikosch, Roberto Rigobon, Jose Scheinkman, Yongzhao Shao, Assaf Zeevi. We also thank participants at NHH, the Norwegian Central Bank, the Columbia Risk Seminar, Universidad Carlos III Madrid, the 2005 International Conference on Finance in Copenhagen, the Federal Reserve Bank of New York, the Federal Reserve Bank of Boston, Wesleyan, Pace and MIT.

JEL Classification: C14, C22, C50 F3, G15

Contents

List of Tables

1 Introduction

This research aims to formalize and quantify the observation that economic activity in general, and financial returns in particular, alternates through periods of turbulence and quiescience.[1] A key distinction that I focus on is the tendency of return correlations to increase during turbulent periods. Understanding this downside comovement is critical, since the very hub of portfolio allocation theory is the assumption that not all returns move together at the same time. Consequently, in the international context, an understanding of this phenomenon could be important in resolving the home bias puzzle. In the domestic context, it could help explain the equity premium puzzle–if stocks display such comovement, then their riskiness relative to bonds may be even larger than presumed. Therefore, quantitative information regarding the importance of turbulence may enable us to glean insight into the underlying structure of financial markets. To phrase the question precisely, I am interested in ”How frequent and how persistent is turbulence?”

I address this question using a dynamic copula approach. The two main reasons for using this approach are as follows. First, use of the copula function allows for robust estimation of dependence.[2] Unlike correlation-based inference, the copula extracts the way in which variables comove, regardless of the scale with which the variables are measured.[3] Second, the use of a mixture model is a convenient compromise between parametric and nonparametric estimation approaches, combining the best features of both. Like the nonparametric approach, a mixture model can, in principle, approximate many distributional shapes and is therefore less restrictive than a single, fully parametric model. Like the parametric approach, the mixture model approach allows one to keep the dimension of the parameter space small.[4]

In this paper I estimate a mixture of normal copulas, characterized by different correlation structures. In related research (Hu (2004) and Rodriguez (2004)) the mixed copula approach is also used. However, those papers utilize normal and asymmetric distributions in the mixture. This has the advantage of extracting different tail dependence measures,

[1]See Ang and Chen (2002), Bollerslev, Chou, and Kroner (1992), Longin and Solnik (2001) and Schwert (2002).

[2]See Dias and Embrechts (2004), De la Pena, Ibragimov, and Sharakhmetov (2003), and Embrechts, McNeil, and Straumann (2001)

[3]Furthermore, in the estimation I avoid mis-specification of the marginal distribution by computing the rank-based empirical distributions, before estimating the copula.

[4]Moreover, the dynamic model employed herein is akin to an adaptive mixture model, which has attractive approximation properties. See McLachlan and Peel (2000) for a discussion of mixture models.

that is, to assess the relative mass of joint observations in the left tail of the return distribution. In the current paper, by contrast, the use of only normal copulas delivers a flexible copula, while preserving a familiar framework. In particular, the current paper uses one dependence concept, the familiar linear correlation. The only intuitive leap is that the correlation is indexed by different copulas which reflect, alternatively, turbulent and quiescent markets.[5] My framework therefore retains normality in a robust way through the copula, while permitting some generality via distinct correlations in the mixture densities.

The remainder of the paper is organized as follows. Section 2 details information on the copula. Section 3 motivates the use of the normal dependence structure and dynamic copulas. Sections 4 and 5 discuss the representation and estimation of the unconditional and conditional mixture models, respectively.[6] Section 6 presents the data and results. Section 7 concludes. The Appendix contains more detail on my estimation methodology.

2 Background and Terminology

2.1 Copulas and Dependence

A copula is a distribution function with uniform marginals. In general (Sklar (1959)) any continuous bivariate joint distribution $F_{X,Y}(x,y)$ can be represented by a copula as a function of the marginal distributions, $F_X(x)$ and $F_Y(y)$. That is,

$$F_{X,Y}(x,y) = C(F_X(x), F_Y(y)) = C(u,v), \tag{1}$$

[5] Specifically, I estimate two correlations, corresponding to the two copulas in the mixture. In the initial specification, the two correlations are restricted to start off at different initial values, one larger than the other. Empirically, this restriction is unnecessary, since the correlations end up with very different final values, even when started at similar initial values.

[6] The sections on unconditional mixture models and hidden markov models follow the exposition of McLachlan and Peel (2000), and McLachlan and Krishnan (1997), adapted to accommodate the mixed copula context.

where and $u = F_X(x)$, and $v = F_Y(y)$. Furthermore, application of the chain rule shows that the corresponding density function $f_{X,Y}(x,y)$ can be decomposed as

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y} = \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial x \partial y} = \frac{\partial^2 C(u,v)}{\partial u \partial v} \frac{\partial F_X(x)}{\partial x} \frac{\partial F_Y(y)}{\partial y} \\ &= c(u,v) f_X(x) f_Y(y), \end{aligned} \tag{2}$$

Intuitively, the joint density is the product of the marginal and copula densities.

Why is the above formulation important and useful? One important reason is that it allows researchers to avoid misspecification of the marginal distributions, and to focus directly on the dependence structure. Another reason, particularly relevant for financial research, is that copula-based dependence measures are robust to monotonic transforms. This permits empirical researchers to work with returns or log returns, for example, at their convenience.[7]

3 The Importance of Normality and Dynamic Copulas

3.1 Econometric and Financial Relevance of Normal Copulas

It is valuable to address the issue of functional forms of copulas. A major advantage of using mixed copulas is that it allows one to nest various dependence shapes. In addition, use of mixed copulas also improves model selection, since it permits the data to choose which copula family is most appropriate. However, a potential disadvantage is the increase in the number and type of parameters to estimate. For example, estimating a four component mixed copula would involve at least four dependence parameters and three weights. Moreover, if the copulas are from different families, for example normal and Gumbel, it is not easy to compare the strength of dependence.[8]

[7]By contrast, traditional correlation based measures of dependence are not necessarily robust to monotonic transforms. SeeEmbrechts, McNeil, and Straumann (2001).

[8]For example, the normal copula's dependence is bounded by 1 and -1, while the Gumbel copula's dependence has no upper bound, in some formulations.

An alternative route is to use a mixture of two normal copulas, which I do in this paper. This approach delivers at least three advantages.[9] First, since both copulas are normal, this formulation allows us to work with a single dependence parameter, the familiar correlation. Moreover, since this parameter has the same range in both copulas, it facilitates explicit comparisons of the strength of dependence. This is not possible with copulas from different families.

Second, the econometric relevance of the two component model is parsimony, since it reduces the number of parameters to be estimated. This is also important in terms of required computer power. Third, the financial insight is that this formulation permits discussion of a fundamental, yet little understood, aspect of financial markets. That is, the importance of turbulent and quiescent periods, which correspond to periods of high dependence and low dependence, respectively. In particular, it allows one to address such important questions as, how *frequent* are turbulent periods relative to quiescent periods? To the best of my knowledge, these questions have never been addressed before in the copula context.

3.2 The Importance of Dynamic Copulas

Dependence has cross section and time series properties. Most research on copulas utilize static copulas. However, in the case of financial markets, there is a clear dynamic element. An important question is, how much emphasis do we want to place on dynamics in the dependence parameters versus dynamics in the *entire* dependence structure? In this paper I choose to focus on the entire dependence structure, since this is more general.[10]

In order to incorporate dynamics, I utilize a hidden markov model (HMM), as discussed in Section 5. It is important to realize that the HMM framework has to be applied with care. In the current application, the HMM structure is natural for at least two reasons. First, there is a reason to use normal copulas, because of the value in working with a single, familiar dependence measure. This permits an immediate interpretation of the estimated parameters. Second, the two component model is also natural, given the economic observation that there are two basic types of dependence structures, corresponding to turbulent

[9]Note that use of a mixture of normals prevents one from addressing issues of tail dependence. In the current setting, however, the research question relates to correlations, and for that purpose the normal copula is appropriate.

[10]Moreover, there appear to be differences in the dependence structure (especially in terms of the estimated weights) in different parts of the sample.

and quiescent periods. This justifies the parsimonious structure, which is attractive for separate econometric and computational reasons. Other applications may or may not have this natural suitedness to the HMM structure. It should also be pointed out that the estimation of parameters and standard errors in this context requires significant computation and derivation, as well as theoretical justification, some of which is provided in the Appendix.[11]

The HMM framework delivers two important advantages. First, it allows one to discern the dynamic structure of dependence. Second, it allows one to answer another important question related to turbulence and quiescence, namely how *persistent* are turbulent periods?[12] To preview the results, it appears that turbulence might be more frequent and persistent than quiescence, a finding which may have useful academic and practical implications.

3.3 Alternative Approaches

Two recent papers have also developed dynamic copulas, albeit in a different context. One uses a change-point approach, and the other uses an ARMA based predictive modelling strategy. We discuss each in turn.

Let us discuss the change-point approach. This approach involves testing for a change in the dependence parameters of a fixed copula family, and is utilized by Dias and Embrechts (2004). A simple example could be to test the constancy of the correlation parameters in a family of bivariate normal copulas $C_N(u, v; \rho)$. For a sample of T observations one can test the null hypothesis of structurally constant parameters as

$$H_0 : \rho_1 = \rho_2 = ... = \rho_T$$

versus

$$H_A : \rho_1 = ... = \rho_{t^*} \neq \rho_{t^*+1} = ... = \rho_T.$$

[11]For example, the estimation of parameters is simplified by the decomposition of the density into copula and marginals, where the latter are not included into the likelihood. Furthermore, the provision of standard errors requires application of both knowledge that a mixture of copulas is a copula, and that the asymptotic covariance matrix can be consistently estimated with the incomplete data Fisher Information matrix.

[12]Frequency of turbulent periods is assessed by the size of the weight on the high correlation copula, and persistence is assessed by the size of the markov transition probability.

A rejection of the null hypothesis indicates that a change of parameter occurred at time t^*. The test statistic is based on a likelihood ratio test, whose critical values can be evaluated via simulation.

The second approach is related to the first, and involves a mix of predictive time-series modelling and copula estimation. This approach is used by Patton (2004), who builds a copula model for portfolio choice between large and small cap stocks, allowing the copula parameters to evolve, depending on certain conditioning variables. In particular, he models the dependence parameters in nine different copulas as a function of lagged risk-free rate, default spread, dividend yield and conditional mean forecasts for the different stocks.[13] Patton finds that knowledge of the dependence structure leads to significant gains for unconstrained investors.

Both of the above papers utilize a single underlying copula. Unlike Patton (2004) and Dias and Embrechts (2004), the hidden markov approach that I employ focuses on clustering, and allows the underlying copula to vary from period to period. I will return to this in Section 5.

4 Unconditional Model

In order to develop my benchmark mixed copula model, it is necessary to clarify some background material on mixture models. Therefore, in this section I present a general mixture model, which I specialize to a mixture of copulas. A general g-component mixture model for data $\mathbf{Y}_t$ observed at time t may be expressed as

$$f(\mathbf{y}_t; \mathbf{\Psi}) = \sum_{i=1}^{g} \pi_i f_i(\mathbf{y}_t; \theta_i), \tag{3}$$

[13] Patton (2004) uses copulas from the following families, normal, student-t, Clayton, Gumbel, Joe-Clayton and Placket. Except for the normal and Plackett, each of these families features unique tail dependence properties. This allows Patton (2004) to infer the dependence structure of the data at extremes.

where π_i are the component weights, and $\mathbf{\Psi}$ $=\{\theta_1, ..., \theta_g; \pi_1, ..., \pi_g\}$ is the set of unknown parameters. The likelihood function is $\prod_{t=1}^{T} \sum_{i=1}^{g} \pi_i f_i(\mathbf{y}_t; \theta_i)$, and the log-likelihood is

$$L(\mathbf{\Psi}) = \sum_{t=1}^{T} \ln \sum_{i=1}^{g} \pi_i f_i(\mathbf{y}_t; \theta_i) \tag{4}$$

The main challenge with estimating parameters using (4) is that it involves a log of sums, for which it is often difficult to compute derivatives. Moreover, when we observe data y_t we do not know which component density $f_i()$ generated it. A standard way to solve this is to treat our estimation as a missing data problem, where the missing data comprises the indicator variables $\{z_i\}$. Specifically, we assume that the *complete* data also contains a label $z_{i,t} \equiv (\mathbf{z}_t)_i$, which indicates the relevant component density for each observation y_t,

$$\begin{aligned} z_{i,t} &= 1, \text{if } y_t \in f_i() \\ & \quad 0, \text{ otherwise} \end{aligned} \tag{5}$$

It is assumed in this context that the z_i's are independently and identically distributed, and that the y_t's are independent, conditional on z. Therefore we can write the conditional distribution of f as

$$f(\mathbf{y}_1, ..., \mathbf{y}_T \mid \mathbf{z}_1, ...\mathbf{z}_T; \mathbf{\Theta}) = \prod_{t=1}^{T} \prod_{i=1}^{g} f_i(\mathbf{y}_t; \theta_i)^{z_{i,t}}, \tag{6}$$

where $\mathbf{\Theta}$ denotes the set of distinct, unknown parameters, $\mathbf{\Theta}$ $=\{\theta_1, ..., \theta_g\}$. I will relax the iid assumption in Section 5. Consequently we can define a *complete*-data likelihood as $\prod_{t=1}^{T} \prod_{i=1}^{g} \pi_i^{z_{i,t}} f_i(\mathbf{y}_t; \theta_i)^{z_{i,t}}$, and *complete*-data log-likelihood L_c as

$$L_c(\mathbf{\Psi}) = \sum_{i=1}^{g} \sum_{t=1}^{T} z_{i,t} \ln \pi_i + \sum_{i=1}^{g} \sum_{t=1}^{T} z_{i,t} \ln f_i(\mathbf{y}_t; \theta_i)\}, \tag{7}$$

where the second component does not depend on the weights π_i. For a mixed copula application, one may replace the last term in (7) with the copula density representation in (2). This is detailed below.

What does introduction of the complete-data likelihood achieve? The quick answer is that it allows maximum likelihood estimation of the parameters using the EM algorithm, a well developed methodology from the statistics literature on incomplete data. Another answer is that expressing the log-likelihood as a function of the z_is permits easy introduction of interesting cross-sectional and dynamic behavior, simply by generalizing the z structure. I will exploit this flexibility in Section 5.

4.1 Mixed Copula and 2-Component Example

In my application of the mixture model, I specialize the functions in (3) to include copulas. That is, the densities f_i are copulas, and as such represent the joint dependence structure of the pertinent variables. Other variables in (3) have the standard mixture model interpretation. Using the density decomposition in (2) permits the copula log-likelihood function for (3) to be written in the following form,

$$\begin{aligned} L(\rho|x,y) &= \ln \prod_{i=1}^{n} f(x_i, y_i; \theta) = \sum_{i=1}^{n} \ln f(x_i, y_i; \theta) \\ &= \sum_{i=1}^{n} \ln c(u_i, v_i; \theta) + \sum_{i=1}^{n} \left[\ln f_X(x_i) + \ln f_Y(y_i)\right], \end{aligned} \tag{8}$$

where θ represents the set of parameters in the copula, $\theta = (\theta_1, \theta_2)$. The last term, $\sum_{i=1}^{n} \left[\ln f_x(x_i) + \ln f_y(y_i)\right]$, is not affected by the dependence parameters θ. Therefore, our maximum likelihood estimator is

$$\hat{\theta} = \arg\max L(\rho|x,y) = \arg\max \sum_{i=1}^{n} \ln c(u_i, v_i; \theta) \tag{9}$$

I now show an example using the above approach for a parametric mixed copula. In particular, I utilize two normal copulas, therefore the functional form of the mixed copula is

$$H(x,y) = C_{mix}(u,v) = w_1 C_1(u,v;\rho_1) + w_2 C_2(u,v;\rho_2), \tag{10}$$

where C_1 and C_2 are both normal copulas, indexed by correlation coefficients ρ_1 and ρ_2, respectively. The normal copula and its density are presented in the Appendix. Following the exact steps outlined above, I use the density representation from (2) to obtain the following expression for the mixed copula model,

$$h(x, y) = [w_1 c_1(u, v; \rho_1) + w_2 c_2(u, v; \rho_2)]\, f_X(\Phi^{-1}(u)) f_Y(\Phi^{-1}(v))$$

The corresponding log-likelihood function can therefore be written as

$$\begin{aligned} L_{mix}(\rho) &= \sum_{t=1}^{T} \ln\left[w_1 c_1(u_t, v_t; \rho_1) + w_2 c_2(u_t, v_t; \rho_2)\right] + \\ &\quad \sum_{t=1}^{T} \left[\ln f_X(\Phi^{-1}(u_t)) + \ln f_Y(\Phi^{-1}(v_t))\right] \end{aligned}$$

The second component of the likelihood is irrelevant for the maximization since it does not contain ρ.

5 Dynamic Copula Model

The dependence structure of financial markets is dynamic: firms and regions alternatively flourish or decline, come into existence or die. Consequently, comovement patterns of financial returns are ever-evolving. It is important for a realistic copula model to account for such dynamic behavior.

5.1 Hidden Markov Model

Hidden markov models (HMMs) are extensions of the mixture models described in Section 4. HMMs provide a straightforward means of allowing dynamic behavior in the copula. Formally, a hidden markov model is a set of states, each with a probability distribution. The transition probabilities π_{hi} determine movement between the different states. The true states and transition probabilities are unobservable, and therefore have to be estimated. [14] This adds an extra set of parameters to our unconditional mixture model in the preceding section. Why would we be interested in doing this? The simple answer is that it allows us to be more flexible about the form of dependence. At a more fundamental level, we

[14]The HMM has five components, the number of states (2 in this case), the number of observations (T), the transition probabilities ($\Lambda = \{\pi_{hi}\}$), a probability distribution in each state, and the initial state distribution π_{0i}. Therefore HMMs generalize mixture models to account for state transition dynamics.

would like to predict dependence and risk, for academic as well as practical finance questions. This framework is developed in greater detail in the Appendix, and to the best of my knowledge, this is the first paper to utilize this framework in the mixed copula context. Here I give the most important insights.

Why is the HMM structure appropriate for this enquiry? The reason is that it is a flexible way to allow for time dynamics in multivariate data that is cross-sectionally dependent. This is precisely what we require in this context since our data and estimated parameters may cluster in certain periods. The main distinction in the HMM is that, unlike the unconditional model of Section 4, the z_{it}'s are not independently distributed, because of temporal correlation between observations.[15] Since the weights π_i depend on z_i, the mixture density no longer has the simple form of (3). Therefore estimation of the parameters is less straightforward. In this context, the HMM approach is very useful, since it embeds the dependent $z_{i,t}$'s in the likelihood in a natural way. Specifically, the HMM allows the z's to follow a stationary Markov chain with transition matrix $[\pi_{hi}]$, where $h = 1, ...g$, and $i = 1, ..g$. We can therefore summarize the conditional dynamics of the $z_{i,t}$'s at each period t as

$$\pi_{hi} = pr\{Z_{i,t+1} = 1 \mid Z_{h,t} = 1\}, \tag{11}$$

where the initial distribution is denoted $\pi_{i,0}$,, $i = 1, ...g$.

Estimation of a general HMM is a slightly more involved version of the unconditional mixture model. One can use Newton-type maximum likelihood to estimate the relevant conditional expectation. The standard approach uses a version of the EM algorithm, known as the Baum-Welch algorithm. This procedure augments the E step with forward and backward recursions through the data in order to estimate the transition probabilities. Asymptotic results and computation of standard errors for HMMs are described in the Appendix. What follows is a brief outline of the estimation procedure.

Let the set of initial and transition probabilities be denoted $\Lambda = \{\pi_{i,0}, \pi_{it}\}$, and let $p(z; \Lambda)$ represent the unconditional initial distribution. Using (11) one knows that at time

[15] For example, in the financial time series context, this temporal correlation can exist in the mean (ARMA models) or in the variance (GARCH models) of the observations. I avoid those approaches in this paper since I wish to place no parametric restrictions on the marginal distributions.

t, the $z_{i,t}$'s depend only on $t-1$, and can represent the distribution of Z in the following manner[16],

$$p(z;\Lambda) = \prod_{i=1}^{g} \pi_{i,0}^{z_{i1}} \prod_{t=2}^{T} \prod_{h=1}^{g} \prod_{i=1}^{g} \pi_{hi}^{z_{h,t-1} z_{i,t}}. \tag{12}$$

The likelihood function now contains an additional term reflecting the distribution of Z, and is written as $p(z;\Lambda) \prod_{t=1}^{T} \prod_{i=1}^{g} f_i(\mathbf{y}_t;\theta_i)^{z_{it}}$. The corresponding complete log-likelihood[17] is

$$\begin{aligned} \ln L_c(\Psi) &= \ln p(\mathbf{z}) + \sum_{i=1}^{g} \sum_{t=1}^{T} z_{it} \ln f_i(\mathbf{y}_t;\rho_i) \\ &= \sum_{i=1}^{g} z_{i1} \ln \pi_{i,0} + \sum_{h=1}^{g} \sum_{i=1}^{g} \sum_{t=1}^{T-1} z_{h,t} z_{i,t+1} \ln \pi_{hi} + \\ &\quad \sum_{i=1}^{g} \sum_{t=1}^{T} z_{it} \ln f_i(\mathbf{y}_t;\theta_i) \end{aligned} \tag{13}$$

It is instructive to compare this to the unconditional case of Section 4. The main difference between this HMM log-likelihood and (7) is the additional middle term, reflecting the entire sequence of transition states over time. In Section 4 the model did not require this term, since it was assumed that $z_{i,t}$ was iid and therefore had no relevant information in its dynamic structure.

5.2 Mixed Copula Application

For a mixed copula application, one may replace the last term in (13) with the copula density representation in (2). This yields

$$\begin{aligned} \ln L_c^{copula}(\Psi) &= \sum_{i=1}^{g} z_{i1} \ln \pi_{i,0} + \sum_{h=1}^{g} \sum_{i=1}^{g} \sum_{t=1}^{T-1} z_{h,t} z_{i,t+1} \ln \pi_{hi} + \sum_{i=1}^{g} \sum_{t=1}^{T} z_{it} \ln c_i(u_t, v_t;\theta_i) + \\ &\quad \sum_{i=1}^{g} \sum_{t=1}^{T} z_{it} \ln f_X(F_X^{-1}(u_t)) f_Y(F_Y^{-1}(v_t)). \end{aligned}$$

[16] The first term is the initial probability of z, and is based on the initial distribution of the Markov chain. The subsequent probabilities of z depend on all the transition probabilities for the Markov chain.

[17] In this case, the parameter vector Ψ contains not only the weights and density parameters but also the transition probabilities, $\Psi = \{\Theta, \Lambda\}$.

We can ignore the last term, in the maximization since it contains no parameters relevant to the copula. The log likelihood therefore becomes

$$\ln \bar{L}_c^{copula}(\Psi) = \sum_{i=1}^{g} z_{i1} \ln \pi_{i,0} + \sum_{h=1}^{g} \sum_{i=1}^{g} \sum_{t=1}^{T-1} z_{h,t} z_{i,t+1} \ln \pi_{hi} + \sum_{i=1}^{g} \sum_{t=1}^{T} z_{it} \ln c_i(u_t, v_t; \rho_i).$$

The estimation of the copula HMM is discussed in the Appendix.

5.3 Model Selection

The various copula models are compared using both information criteria and empirical distance measures. I discuss each in turn. First, let us consider the information criteria. The Akaike Information Criterion (AIC) is used because of its optimality properties, and also since it can be used to compare nested or non-nested models, in contrast to tests based on likelihood ratios, for example. The AIC is used in other multivariate model selection research, for example in Burnham and Anderson (2002), Dias and Embrechts (2004), and Rodriguez (2004).[18] We also use the Bayes Information Criterion (BIC). While the BIC does not share the same optimality properties as AIC, it penalizes more strictly for overfitting a model, which is sometimes a desirable property. The standard expressions for AIC and BIC are as follows. Consider a sample with size equal to T, and the number of estimated parameters (θ) equal to q. Then the AIC and BIC are defined as

$$\begin{aligned} AIC(q) &= -2\ln[\hat{L}(\theta)] + 2q \\ BIC(q) &= -2\ln[\hat{L}(\theta)] + q\ln(T). \end{aligned}$$

The best model is selected as the one that minimizes AIC or BIC.

Second, let us consider the empirical distance measures, namely, the Kolmogorov-Smirnov (KS) and Andersen-Darling (AD) distances. These measure the distance be-

[18] If we denote the true density f, then the standard information-based measure for choosing the best approximate model g is the Kullback-Leibler Information (K):

$$K(f,g) = \int f(x) \ln \frac{f(x)}{g(x|\theta)} dx.$$

Mathematically speaking, K(f,g) is a directed distance from candidate models to the true model. The AIC (and its small sample adjusted counterparts) is generally an unbiased estimator of K, hence its desirability in model selection. For more discussion of these considerations, see Burnham and Anderson (2002), Chapter 2.

tween the empirical distribution and the estimated distribution corresponding to the maximized likelihood. Specifically, consider a data set $\{X_1, ...X_T\}$ with associated parameters $\theta = \{\theta_1, ..., \theta_T\}$, and a parametrically estimated distribution function $\hat{F}(\hat{\theta})$. If we write the empirical distribution function as $F_T(\theta)$, then the Kolmogorov-Smirnov distance KS may be expressed as

$$KS = \max_X \mid \hat{F}(\hat{\theta}) - F_T(\theta) \mid,$$

and the Anderson-Darling distance AD is

$$AD = \max_X \frac{\mid \hat{F}(\hat{\theta}) - F_T(\theta) \mid}{\sqrt{F_T(\theta) * [1 - F_T(\theta)]}}.$$

The average KS and AD are computed by integrating with respect to the empirical density. In the current application, with uniform densities, the integral reduces to a simple arithmetic mean.

6 Data and Results

The data used in this paper comprises weekly observations on equity indices for five countries, France, Germany, Japan, the United Kingdom, and the United States. These data are available from the MSCI data base. The sample period is from 1/20/90 to 5/29/02, for a total of 646 observations.

The results from the estimation methodology of Sections 4 and 5 are presented in Tables 1 through 7. In addition to estimated parameters, I also present the Akaike and Bayesian Information Criteria (AIC and BIC) for assessing goodness of fit. As mentioned previously, the unconditional model has two component densities, the ith component being a normal copula with correlation ρ_i. Standard errors are computed using the methods developed in the Appendix. I now discuss each table, in turn.

Table 1 reports estimates of the unconditional model. Since the research question concerns the relative importance of high- and low-correlation regimes, I focus on the parameters w_1 and w_1. These parameters represent the frequency of low and high correlation regimes, respectively. In all cases except for the France-Japan and Germany–Japan pairs, the greater weight is on the component with the larger correlation. That is, w_2 exceeds w_1 in all but two country pairs. The largest correlation is 0.826, for ρ_2 in the France-Germany pair. The smallest (absolute value) correlation is -0.028, for ρ_1 in the Japan-UK pair. The largest spread between ρ_1 and ρ_2 is for the UK-US pair, a value of 0.758. This indicates that there is a huge difference between the correlation during normal and turbulent times.

Tables 2 and 3 report results from dividing the weekly data in two equal-sized samples. Table 2's results are for the first half of the sample. The evidence on weights is weaker in this sample. Specifically, greater weight is on the component with the larger correlation in six of the pairs, all except France-Japan, Germany-Japan, Germany-US and Japan-UK. The largest correlation is 0.862, for ρ_2 in the Japan-UK pair. The smallest correlation is -0.055, for ρ_1 in the Germany-Japan pair. The largest spread between ρ_1 and ρ_2 is for the France-Germany pair, a value of 1.177.

Table 3 presents results from the second half of the weekly sample. These results are more similar to those for the full sample. In all pairs except for France-Japan and France-US the greater weight is on the component with the larger correlation. That is, w_2 exceeds w_1 for all but two country pairs. The largest correlation is 0.905, for ρ_2 in the France-US pair. The smallest correlation is -0.129, for ρ_1 in the Germany-Japan pair. The largest spread between ρ_1 and ρ_2 is for the Japan-UK pair, 1.334. Taken together, the results in Tables 2 and 3 indicate that there may be some interesting dynamic behavior in the data. This dynamic behavior is explored in the hidden markov model, whose results I now discuss.

Table 4 presents results for the hidden markov model. Since the research question relates to persistence of high-correlation regimes, I focus on the parameters π_{11} and $\pi_{22}\dot{}$These parameters represent the probability of remaining in a low- and high-correlation regime, respectively. The most striking finding is that in all country pairs except France-Japan, π_{22} exceeds 1/2. That is, the likelihood of staying in a high-correlation regime is very compelling. However, π_{22} is not always the largest transition probability. The largest correlation is 0.836, which is ρ_2 for the France-Germany pair. The smallest correlation is ρ_1 equal to -0.076, for the Japan-UK pair, and the largest spread is 0.776, for the Japan-US pair. Moreover, the AIC and BIC indicate that the hidden markov model fits much better for each country pair, relative to the unconditional model in Table 1.

Table 5 demonstrates additional estimation results from a hidden markov model comprising a normal and rotated gumbel mixture. I denote this model the ”normal-RG model” for brevity. This mixture model is a reasonable alternative to the double normal hidden markov model, since it can represent financial returns that are subject to alternating periods of normality and downside risk.[19] The dependence parameters are all significantly different from zero. It is not, however, easy to compare the dependence parameters, since they belong to different copula families. The largest transition probabilities are π_{11} or π_{22} in eight of the ten country pairs, with the exception of Japan-UK and UK-US. This indicates that there is persistence in both periods of normality and periods of downside risk.

Table 6 presents results of the Kolmogorov-Smirnov and Anderson-Darling statistics, as well as their corresponding averages. The models I compare include the 2-component normal copula, No(2), the mixture of normal and Rotated Gumbel copulas, NoRG(2), and the best of a set of other unconditional copulas, BestUnc. [20] It should be noted that the hidden markov models are not included in this table because they do not indicate a specific density for each observation, and therefore cannot be directly compared with the empirical density. In all cases the best fit comes from the ”best unconditional model”, which includes a set of single and mixed *non-normal* copulas. In other words, the mixed normal copulas from Section 4 are dominated by other unconditional copulas.

Table 7 shows a comprehensive comparison of the various copula models, based on the AIC and BIC, in Panels A and B, respectively. We discuss each, in turn. Panel A presents the AIC results. The striking result is that the hidden markov models consistently dominate all other models, with a rank of 1 or 2 for each country pair. The normal hidden markov model does the best, with an average rank of 1.2. The next best model is the normal-RG hidden markov model, which in most cases is very close to the normal hidden markov model The poorest performance is the single normal copula, with an average rank of 4.6.

Panel B displays the BIC results. The best models are again overwhelmingly the normal and normal-RG hidden markov models, with ranks of 1.2 and 1.9, respectively. The poorest performance is the unconditional double normal copula, with an average rank of 4.4. In sum, the two main findings in Table 7 are as follows. First, the double normal copula improves on a single normal copula according to the AIC but not according to the BIC, which favors parsimony. Second, and more strikingly, the hidden markov models overwhelmingly dominate other models, including the non-normal models (Best unconditional) that outperformed the *unconditional* normal mixtures in Table 6. This dominance of the HMM suggests that a dynamic normal copula can capture some aspects of dependence

[19] This intuitive formulation and interpretation was suggested by Bob Hodrick.

[20] To clarify further, BestUnc refers to the best fitting copula from a broad set of Archimedean, extreme-value, and elliptical copulas. These include the families of Frank, student-t, and Gumbel.

behavior better than copulas with explicit tail dependence.[21] It should of course be borne in mind that a mixture of normal copulas will have limited ability to capture asymptotic tail behavior.[22] Therefore the choice of which copula formulation to use will depend heavily on the application of interest.

7 Conclusions

In this research, I investigate the structure of dependence in financial markets using a mixed copula approach. This methodology is closely related to the regime-switching methodology in the time series literature. [23] This paper presents three technical contributions, two conceptual contributions, and three empirical contributions. The first technical contribution is the introduction of an econometrically parsimonious dependence structure, namely, a mixture of normal copulas, that still delivers valuable economic insights. The second technical contribution is the implementation of a hidden markov structure to account for time variation in the structure of dependence. Third, in the Appendix I extend previous results to obtain standard errors for dynamic mixed copulas. It is important to note that, within the relevant mixture component, the correlation is a robust dependence measure, since the component is a copula and therefore rank-based. In this way, I use a familiar correlation measure, but ”robustify” it by allowing it to come from one of several copulas at any point in time, and allow for persistence in the choice of copula.

The two conceptual contributions are first, relating the copula technology to characterization of turbulent and quiescent periods, and second, addressing the specific questions, how frequent and how persistent are turbulent periods? The three most striking empirical findings are as follows. First, for the unconditional model, the weights are generally greater for the copula indexed with a large correlation. This indicates the interesting result that turbulent periods are more common. Moreover, for the unconditional model, the weights and parameter estimates differ greatly between the two sub-samples. Second, the hidden markov model provides a much better fit than the unconditional model. Third,

[21] This is consistent with evidence on improved goodness of fit for dynamic copulas, in Dias and Embrechts (2004).

[22] It can be shown that at least one component density must have tail dependence for an unconditional mixture to have tail dependence. This holds for stationary dynamic models, where for example, we calculate the weight π_1 from its stationary distribution as $\pi_1 = (1-\pi_{22})/(2-\pi_{11}-\pi_{22})$. I am grateful for discussions with Jonas Andersson, Jostein Lillestol and Ching-Chih Lu on this point.

[23] See, for example, Chapter 22 of Hamilton (1994).

the estimated transition probabilities for the dynamic copula indicate turbulent periods are often persistent.

Why should we care about these findings? There are at least two important reasons. First, when taken together, these results indicate that international financial markets might be prone to episodes of persistent instability. This possibility has clear implications for theory and practice of finance, which typically assume the generic existence of stable economies. Second, on a practical level, a major implication of this paper is that the gaussian assumption might still be utilized in financial modelling and asset allocation, once one generalizes in a suitable manner for correlation clustering and dynamic behavior. Future work could extend this framework to account for more general dynamics, both in theoretical and empirical applications, and formalize the links to research on financial contagion.

References

Ang, Andrew, and Joseph Chen, 2002, Asymmetric Correlations of Equity Portfolios, *Journal of Financial Economics* 63, 443–94.

Bickel, Peter, Yakov Ritov, and Tobias Ryden, 1998, Asymptotic Normality of the Maximum-Likelihood Estimator for General Hidden Markov Models, *The Annals of Statistics* 26, 1614–1635.

Bollerslev, Tim, Ray Chou, and Kenneth Kroner, 1992, ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics* 52, 5–59.

Burnham, Kenneth P., and David R. Anderson, 2002, *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach, 2nd Edition*. (Springer-Verlag New York).

De la Pena, Victor, Rustam Ibragimov, and Shaturgun Sharakhmetov, 2003, Characterizations of Joint Distributions, Copulas, Information, Dependence and Decoupling, Working paper, Columbia University.

Dempster, A. P., N. M. Laird, and Donald B. Rubin, 1977, Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)* 39, 1–38.

Dias, Alexandra, and Paul Embrechts, 2004, Dynamic copula models for multivariate high-frequency data in finance, Working paper, .

Embrechts, Paul, Alexander McNeil, and Daniel Straumann, 2001, Correlation and Dependence in Risk Management: Properties and Pitfalls, Working paper, Departement Mathematik, ETHZ.

Hamilton, James D., 1994, *Time Series Analysis*. (Princeton Press New Jersey).

Hu, Ling, 2004, Dependence Patterns across Financial Markets: a Mixed Copula Approach, Working paper, Department of Economics, The Ohio State University.

Longin, Francois, and Bruno Solnik, 2001, Extreme Correlation of International Equity Markets, *Journal of Finance* 56, 649–76.

Louis, T., 1982, Finding the Observed Information Matrix when using the EM Algorithm, *Journal of the Royal Statistical Society, Series B* 44, 226–233.

McLachlan, Geoffrey J., and Thriyambakam Krishnan, 1997, *The EM Algorithm and Extensions*. (John Wiley & Sons, New York).

McLachlan, Geoffrey J., and David Peel, 2000, *Finite Mixture Models*. (John Wiley & Sons, New York).

Patton, Andrew J., 2004, On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation, *Journal of Financial Econometrics* 2, 130–168.

Rodriguez, Juan Carlos, 2004, Measuring Financial Contagion: A Copula Approach, Working paper, EURANDOM.

Schwert, G. William, 2002, Stock Volatility in the New Millenium: How wacky is Nasdaq?, *Journal of Monetary Economics* 49, 3–26.

Sklar, Abraham, 1959, Fonctions de repartition a n dimensions et leurs marges, *Pub. Inst. Statist. Univ. Paris* 8, 229–231.

8 Appendix

8.1 Distribution and Density of the Normal Copula

Let $\Phi_\rho(x,y)$ denote the standard bivariate normal cumulative distribution. In other words,

$$\Phi_\rho(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\pi \mid \Sigma \mid} \exp\{-\frac{1}{2}(xy)\Sigma^{-1}(xy)'\}dxdy,$$

where $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. From this one may obtain the following distribution function for the normal copula, denoted $C_N(.)$:

$$C_N(u,v;\rho) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)).$$

The corresponding copula density for the preceding copula distribution is represented below, and denoted $c_\rho(.)$. In both the density and distribution case, the dependence parameter ρ lies between -1 and $+1$.

$$c_\rho(.) = \frac{1}{\sqrt{1-\rho^2}} \exp\left[\frac{-[\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2 - 2\rho\Phi^{-1}(u)\Phi^{-1}(v)]}{2(1-\rho^2)} + \frac{\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2}{2}\right].$$

8.2 Estimation of Unconditional Copula Model

I now return to the general formulation in (7). The estimation method for parameters in (7) is based on the EM algorithm of Dempster, Laird, and Rubin (1977). This procedure comprises two steps. At each iteration k, the first (E) step takes the conditional expectation of (7), $Q(\mathbf{\Psi}^{(k)})$, given the data. Then the second (M) step maximizes $Q(\mathbf{\Psi}^{(k)})$ to obtain the updated parameter estimates, which are substituted into the E step to obtain updated estimates of the posterior probabilities[24]. This iterative procedure continues until the estimated log-likelihood function reaches a maximum. The details are now discussed.

[24] The posterior probability $\tau_i(y_t;\Psi^{(k)})$ denotes the expectation that observation y_t belongs to the ith component density. Intuitively, the posterior probabilities are updated every iteration using Bayes' rule.

8.2.1 E-Step

The initial iteration $(k = 0)$ takes the conditional expectation $Q(\mathbf{\Psi}^{(0)})$ using the initial parameter values $\mathbf{\Psi}^{(0)}$. Then, on the $(k+1)$th iteration, one computes the posterior probabilities $\tau_i(\mathbf{y}_t; \mathbf{\Psi}^{(k)})$ as

$$\tau_i(\mathbf{y}_t; \mathbf{\Psi}^{(k)}) = \frac{\pi_i^{(k)} f_i(\mathbf{y}_t; \theta_i^{(k)})}{\sum\limits_{h=1}^{g} \pi_h^{(k)} f_h(\mathbf{y}_t; \theta_h^{(k)})}, \tag{14}$$

which are used to compute the conditional expectation of the log-likelihood,

$$Q(\mathbf{\Psi}^{(k)}) = \sum_{i=1}^{g} \sum_{t=1}^{T} \tau_i(\mathbf{y}_t; \mathbf{\Psi}^{(k)}) \{\ln \pi_i + \ln f_i(\mathbf{y}_t; \theta_i)\}. \tag{15}$$

8.2.2 M-Step

The M-step chooses parameters $\mathbf{\Psi}^{(k+1)}$ to maximize (15), which yields an updated set of parameters. Specifically, on iteration $k+1$ one chooses $\mathbf{\Psi}^{(k+1)}$ as

$$\mathbf{\Psi}^{(k+1)} = \arg\max \sum_{i=1}^{g} \sum_{t=1}^{T} \tau_i(\mathbf{y}_t; \mathbf{\Psi}^{(k)}) \{\ln \pi_i + \ln f_i(\mathbf{y}_t; \theta_i)\}. \tag{16}$$

It should be noted that the weights are calculated as the mean of the posterior probabilities, that is, $\pi_i^{(k+1)} = \frac{1}{T} \sum\limits_{t=1}^{T} \tau_i(\mathbf{y}_t; \mathbf{\Psi}^{(k)})$. This is intuitive, since the estimated weight on component density i reflects the best current estimate of being in i.

8.3 Estimation of Dynamic Copula Model

I now describe the estimation of a general HMM, derived from a copula mixture model.[25] As in the unconditional mixture model, we use Newton-type maximum likelihood to estimate the conditional expectation of (13), $\tilde{Q}()$. The standard approach uses a version of the EM algorithm, known as the

[25] I feel this explication is necessary since this is the first use of the HMM approach in this context–there seems to be no other literature that estimates dynamic dependence using copulas and the hidden markov framework. Although the derivation is the same as that in some of the HMM literature, e.g. Hamilton (1994) and McLachlan and Peel (2000), it is necessary to represent all the details, since this permits easy and clear understanding of the steps in the mixed copula application. Moreover, this development allows other researchers involved in dynamic dependence to have a concrete starting point with detailed steps. Readers familiar with the estimation of hidden markov mixture models may proceed directly to the next section.

Baum-Welch algorithm. This procedure augments the E step with forward and backward recursions through the data in order to estimate the transition probabilities.[26]

8.3.1 E-Step

In this step one calculates $\tilde{Q}()$, the conditional expectation of (13) from the observed sample. One iterates on the expectation recursively until the improvements in the log-likelihood function fall below our criterion. Because of the middle term in (13), the E-step itself has three steps, defining the probabilities, computing auxiliary probabilities, and substituting into the $\tilde{Q}()$ function.

The first step, one sets up the conditional probabilities and define forward and backward probabilities. Let $\tau_{hi,t}^{(k)}$ and $\tau_{i,t}^{(k)}$ represent the following conditional probabilities at time t,

$$\tau_{hi,t} = pr\{Z_{h,t} = 1, Z_{i,t+1} = 1 \mid \mathbf{y}\} \tag{17}$$

and

$$\tau_{i,t} = pr\{Z_{i,t} = 1 \mid \mathbf{y}\}.$$

Furthermore, one can calculate the probability of being in state i today as the sum of all yesterday's probabilities of moving to state i, from any state, $\tau_{i,t} = \sum_{h=1}^{g} \tau_{hi,t-1}$. The initial probability of being in state i, $\tau_{i,1}$ is estimated using the initial values for the Markov chain, $\tau_{i,1} = \pi_{i,0} f_i(\mathbf{y}_1) / \sum_{h=1}^{g} \pi_{h,0} f_h(\mathbf{y}_1)$.

However, the values of $\tau_{hi,t}$ are still unknown. In order to obtain them, re-express (17) using Bayes' rule,

$$\tau_{hi,t} = pr\{Z_{h,t} = 1, Z_{i,t+1} = 1 \mid \mathbf{y}\} = \frac{pr\{Z_{h,t} = 1, Z_{i,t+1} = 1\}}{pr\{\mathbf{Y} = \mathbf{y}\}} \tag{18}$$

This simplifies the expression for $\tau_{hi,t}$ because one can evaluate the ratio on the right hand side in (18) using two auxiliary probabilities. These variables are the "forward" probabilities $a_{i,t}$ and the "backward" probabilities $b_{i,t}$, defined in the following manner,

$$\begin{aligned} a_{i,t} &= pr\{Y_1 = y_1, ..., Y_t = y_t,\ Z_{it} = 1\}, \quad t = 1, ..., T \\ b_{i,t} &= pr\{Y_{t+1} = y_{t+1}, ..., Y_T = y_T \mid Z_{it} = 1\}, \quad t = T-1, T-2..., 1 \end{aligned} \tag{19}$$

[26] The addition of forward and backward recursions is more efficient for computing the middle term in (13) than summing over all possible state sequences from $t = 1 : T$, since it calculates and stores results for later use, instead of recomputing at each node. The logic is based on dynamic programming, as discussed in Rabiner (1989).

Combining the (19) and (18) one obtains the following equation for $\tau_{hi,t}$,

$$\tau_{hi,t} = \frac{a_{h,t}\pi_{hi}f_i(y_{t+1})b_{i,t+1}}{\sum_{h=1}^{g}\sum_{i=1}^{g} a_{h,t}\pi_{hi}f_i(y_{t+1})b_{i,t+1}} \tag{20}$$

The second step involves computation of (20). To do this, forward and backward recursions are computed at each iteration to estimate values for $a_{i,t}^{(k)}$ and $b_{i,t}^{(k)}$, which are then substituted in (20) before proceeding to the M step.

Forward Recursions The forward recursions consist of two steps, induction and termination. The initial value is set to $a_{i,t}^{(k)} = \pi_{0,i}^{(k)} f_i^{(k)}(\mathbf{y}_1), \quad i = 1, ..., g$. The induction step at iteration $k+1$ involves searching forward through the data from period 1 onwards. Specifically, at iteration $k+1$, the forward probability $a_{i,t+1}$ is computed as

$$a_{i,t+1}^{(k)} = [\sum_{h=1}^{g} a_{h,t}^{(k)}\pi_{hi}^{(k)}]f_i^{(k)}(\mathbf{y}_{t+1}), \quad t = 1, ..., T-1.$$

The termination step is determined as

$$pr_{\Psi^{(k)}}(Y_1 = y_1, ..., Y_T = y_T,) = \sum_{i=1}^{g} a_{it}^{(k)}.$$ [27]

Backward Recursions The backward recursion has only an induction step. The initial value is set to unity, $b_{h,T}^{(k)} = 1, \ \ h = 1, ..., g$. The induction step at iteration $k+1$ involves searching *backward* through the data starting at period $T-1$. Specifically,

$$b_{h,t}^{(k)} = \sum_{i=1}^{g} \pi_{hi}^{(k)} f_i^{(k)}(\mathbf{y}_{t+1})b_{i,t+1}^{(k)}, \quad t = T-1, ..., 1; \ \ h = 1, ..., g.$$

As mentioned before, the estimated $b_{i,t}$s and $a_{i,t}$s are substituted in (20) before proceeding to the M step. Specifically, on the $k-$th iteration one computes $\tau_{hi,t}$ as

$$\tau_{hi,t}^{(k)} = \frac{a_{h,t}^{(k)}\pi_{hi}^{(k)}f_i^{(k)}(y_{t+1})b_{i,t+1}^{(k)}}{\sum_{h=1}^{g}\sum_{i=1}^{g} a_{h,t}^{(k)}\pi_{hi}^{(k)}f_i^{(k)}(y_{t+1})b_{i,t+1}^{(k)}}, \ t = 1, ..., T-1$$

[27] The operator $\mathrm{pr}_{\Psi^{(k)}}$ denotes probability conditional on information available at the kth iteration.

In the third step, one estimates $\tilde{Q}(\mathbf{\Psi}^{(k)})$ on iteration $k+1$ as

$$\tilde{Q}(\mathbf{\Psi}^{(k)}) = \sum_{i=1}^{g} \tau_{i,1}^{(k)} \ln \pi_{i,0} + \sum_{h=1}^{g} \sum_{i=1}^{g} \sum_{t=1}^{T-1} \tau_{hi,t}^{(k)} \ln \pi_{hi} + \sum_{i=1}^{g} \sum_{t=1}^{T} \tau_{i,t}^{(k)} \ln f_i(\mathbf{y}_t; \theta_i). \tag{21}$$

8.3.2 M-Step

In the M step the updated estimates are computed using (21). In particular, on iteration $k+1$ one has

$$\pi_{0,i}^{(k+1)} = \tau_{i,1}^{(k)},$$

and

$$\pi_{hi}^{(k+1)} = \frac{\sum_{m=1}^{T-1} \tau_{hi,t}^{(k)}}{\sum_{t=1}^{T-1} \tau_{i,t}^{(k)}}.$$

At this stage the parameter estimates $\hat{\theta}_i$ are obtained recursively using Newton-type maximum likelihood approach.[28]

8.4 Standard Errors for Dynamic Copula Models

In this subsection I present and build on existing results for HMM standard errors.[29] Unlike previous research, we apply this methodology to the case of mixed copulas. For clarity, I express the computational aspects of the standard errors since these are utilized in my empirical applications. The paper of Bickel, Ritov, and Ryden (1998) establishes consistency and asymptotic normality of the MLEs, $\hat{\alpha}_T$, for HMMs, and shows that the asymptotic covariance matrix is estimated consistently by the observed Fisher information, $-\frac{1}{T}\frac{\partial^2}{\partial\alpha\partial\alpha'}L(\hat{\alpha}_T)$. The main two results in that research obtain under standard regularity conditions, and are described in their Lemma 2 and Theorem 1 below.

Lemma 2. *Let $\hat{\alpha}_T$ be any sequence in Θ such that $\lim_{T\to\infty} \hat{\alpha}_T = \alpha^*$, almost surely. Then*

$$\frac{1}{T}\frac{\partial^2}{\partial\alpha\partial\alpha'}L(\hat{\alpha}_T) \to -\dot{F} \text{ in } P_0 - \textit{probability as } T \to \infty \tag{22}$$

[28]This is done applying a modification of the EM algorithm, known as the conditional EM (ECM) algorithm. See McLachlan and Krishnan (1997).

[29]It should be noted that the standard error computations presented here are in some cases lower bounds, in instances where there is a large divergence between the complete and incomplete information.

This result establishes consistency of the Hessian of the log likelihood. The next result establishes asymptotic normality of the MLE, where the asymptotic variance is the inverse Fisher information matrix.

Theorem 1. *Assume that the information matrix is nonsingular. Then*

$$\sqrt{T}(\hat{\alpha}_T - \alpha^*) \to^d N(0, \dot{F}^{-1}) \tag{23}$$

In order to obtain standard errors for our mixed copula application, we may implement equation (23), for which it is necessary to compute the Fisher information from the likelihood function. Let us now discuss how to compute the Fisher information for a mixed copula.

As mentioned in Section 4, the likelihood function for a two component mixed copula is

$$L() = p(z; \Lambda) \prod_{t=1}^{T} \prod_{i=1}^{2} f_i(\mathbf{y}_t; \theta_i)^{z_{it}}.$$

Let the likelihood function be $L(\Psi) = \prod f(y_t; \Psi)$, and the score function be $S(Y; \Psi) = \partial \ln L(\Psi)/\partial\Psi$. In general, the Fisher information is

$$I(\Psi) = E_{\Psi}\{S(Y; \Psi)S'(Y; \Psi)\}.$$

Under regularity conditions, the Fisher information can be expressed as

$$I(\Psi; y) = E\left[\frac{-\partial^2 \ln L(\Psi)}{\partial\Psi\partial\Psi'}\right], \tag{24}$$

where the ratio on the right hand side is the negative of the Hessian of the log likelihood function.

The challenge in implementing the Fisher information is that the log likelihood involves missing data. In order to overcome this, it is possible to use an approach suggested by Louis (1982), which we adapt in the following.[30] Note that this approach was originally used to obtain the information matrix for a mixture model. In the current context, we adapt it to the case of a hidden markov model, where the components are mixed copulas. Furthermore, we utilize (23), which establishes that using the Fisher information is justified in the present context.

We now extend the above results and outline how to calculate the asymptotic variance for a mixed copula application. We use the complete data gradient $S(\mathbf{X}, \Psi)$ to compute the observed information. The likelihood for our mixed copula model is as follows:

[30] I am grateful to Ching-Chih Lu for discussions on the following results.

$$\begin{aligned}\ln L_c(\Psi) &= \ln p(\mathbf{z}) + \sum_{i=1}^{2}\sum_{t=1}^{T} z_{it}\ln c_i(\mathbf{y}_t;\rho_i)\\ &= \sum_{i=1}^{2} z_{i1}\ln \pi_{i,0} + \sum_{h=1}^{2}\sum_{i=1}^{2}\sum_{t=1}^{T-1} z_{h,t}z_{i,t+1}\ln \pi_{hi} + \sum_{i=1}^{2}\sum_{t=1}^{T} z_{it}\ln c_i(\mathbf{y}_t;\rho_i)\end{aligned}$$

where $\mathbf{X}$ includes observed data $\mathbf{y}$ and missing data $\mathbf{z}$. Therefore the gradient vector can be represented as

$$S(X_t;\Psi) = \begin{pmatrix} \frac{\partial \ln L_c(\Psi)}{\partial \theta_i} \\ \frac{\partial \ln L_c(\Psi)}{\partial \pi_{hi}} \end{pmatrix} = \begin{pmatrix} z_{it}\frac{\partial \ln c_i(y_t;\theta_i)}{\partial \theta_i} \\ \frac{z_{h,t}z_{i,t+1}}{\pi_{hi}} \end{pmatrix}. \tag{25}$$

We can define the matrix $B(.)$ as

$$B(X_t,\Psi) = \begin{pmatrix} -\frac{\partial^2 \ln L_c(\Psi)}{\partial \theta_i^2} & 0 \\ 0 & -\frac{\partial^2 \ln L_c(\Psi)}{\partial \pi_{hi}^2} \end{pmatrix}, \tag{26}$$

where

$$-\frac{\partial^2 \ln L_c(\Psi)}{\partial \rho_i^2} = -z_{it}\frac{\partial^2 \ln c_i(y_t;\rho_i)}{\partial \rho_i^2} \tag{27}$$

is a 2-by-2 matrix with zeros as off-diagonal elements and $-z_{it}\frac{\partial^2 \ln c_i(y_t;\rho_i)}{\partial \rho_i^2}$ as diagonals, $i = 1,2$. The second derivative matrix is

$$-\frac{\partial^2 \ln L_c(\Psi)}{\partial \pi_{hi}^2} = \frac{z_{h,t}z_{i,t+1}}{\pi_{hi}^2}, \tag{28}$$

which is a 2-by-2 matrix with zeros as all off-diagonal elements.

Since $\pi_{11}+\pi_{12}=1$, $\pi_{21}+\pi_{22}=1$, we have two free parameters. Therefore, the bottom half of equation(25) may be rewriten as

$$\frac{z_{h,t}z_{i,t+1}}{\pi_{hi}} = \begin{pmatrix} \frac{z_{1,t}z_{1,t+1}}{\pi_{11}} - \frac{z_{1,t}z_{2,t+1}}{1-\pi_{11}} \\ \frac{z_{2,t}z_{1,t+1}}{\pi_{21}} - \frac{z_{2,t}z_{2,t+1}}{1-\pi_{21}} \end{pmatrix}. \tag{29}$$

Consequently, equation(28) may be re-expressed as

$$\begin{pmatrix} -\frac{\partial^2 \ln L_c(\Psi)}{\partial \pi_{11}^2} \\ -\frac{\partial^2 \ln L_c(\Psi)}{\partial \pi_{21}^2} \end{pmatrix} = \begin{pmatrix} \frac{z_{1,t}z_{1,t+1}}{\pi_{11}^2} + \frac{z_{1,t}z_{2,t+1}}{(1-\pi_{11})^2} \\ \frac{z_{2,t}z_{1,t+1}}{\pi_{21}^2} + \frac{z_{2,t}z_{2,t+1}}{(1-\pi_{21})^2} \end{pmatrix}. \tag{30}$$

It follows that

$$S^*(y_t, \Psi) = \begin{pmatrix} \tau_{it} \frac{\partial \ln f_i(y_t;\theta_i)}{\partial \theta_i} \\ \frac{\tau_{hi,t}}{\pi_{hi}} \end{pmatrix} \tag{31}$$

$$S^*(\mathbf{y}, \Psi) = E_\theta[S(X, \Psi)|X \in R] \tag{32}$$

where $R = \{x : y(x) = y\}$. Then the complete data information matrix can be expressed as

$$I_Y(\Psi) = E_\theta[B(X, \Psi)|X \in R] - E_\theta[S(X, \Psi)S^T(X, \Psi)|X \in R] + S^*(y, \Psi)S^{*T}(y, \Psi). \tag{33}$$

The result in (33) can be substituted into the asymptotic equation (23) in order to compute the asymptotic covariance matrix. That is, one estimates $\dot{F}$ with $I_Y(\Psi)$.

Table 1: Unconditional Model

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US
$\rho 1$	0.446	0.186	0.371	0.445	0.111	0.098	0.428	-0.028	-0.211	-0.162
	(0.090)	(0.055)	(0.078)	(0.153)	(0.061)	(0.137)	(0.128)	(0.088)	(0.111)	(0.265)
$\rho 2$	0.826	0.799	0.792	0.583	0.754	0.751	0.607	0.632	0.544	0.596
	(0.018)	(0.058)	(0.023)	(0.109)	(0.055)	(0.024)	(0.083)	(0.045)	(0.048)	(0.030)
w1	0.193	0.739	0.298	0.429	0.657	0.205	0.405	0.408	0.341	0.075
w2	0.807	0.261	0.702	0.571	0.343	0.795	0.595	0.592	0.659	0.925
L	276.940	45.642	194.750	101.360	44.614	165.120	106.550	51.571	35.404	110.980
AIC	-547.880	-85.284	-383.500	-196.720	-83.228	-324.240	-207.100	-97.142	-64.808	-215.960
BIC	-534.468	-71.872	-370.088	-183.308	-69.816	-310.828	-193.688	-83.730	-51.396	-202.548

Table 1 presents parameter estimates from the two-component mixed copula. Standard errors are in parentheses. The data frequency is weekly. The sample period is January 10, 1990 to May 29, 2002, containing 646 observations. $\rho 1$,$\rho 2$ and w1, w2 are the correlation parameters and weights in copula 1 and 2, respectively. L represents the log likelihood function evaluated at the estimated parameters, while AIC and BIC are as defined in the paper.

Table 2: Unconditional Model, Sample A

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US
$\rho 1$	-0.420	0.121	0.424	0.379	-0.055	-0.167	0.260	0.282	-0.220	-0.154
	(0.459)	(0.099)	(0.080)	(0.322)	(0.122)	(0.210)	(0.104)	(0.067)	(0.155)	(0.187)
$\rho 2$	0.757	0.704	0.825	0.488	0.690	0.745	0.668	0.862	0.547	0.649
	(0.029)	(0.092)	(0.035)	(0.235)	(0.076)	(0.034)	(0.132)	(0.059)	(0.080)	(0.050)
w1	0.066	0.622	0.447	0.418	0.536	0.219	0.657	0.768	0.406	0.253
w2	0.934	0.378	0.553	0.582	0.464	0.781	0.343	0.232	0.594	0.747
L	109.180	22.684	89.407	33.726	19.966	68.959	28.530	33.091	13.923	40.612
AIC	-212.360	-39.368	-172.814	-61.452	-33.932	-131.918	-51.060	-60.182	-21.846	-75.224
BIC	-200.945	-27.953	-161.399	-50.037	-22.517	-120.503	-39.645	-48.767	-10.431	-63.809

Table 2 presents parameter estimates from the two-component mixed copula. Standard errors are in parentheses. The data frequency is weekly. The sample period is January 10, 1990 to March 20, 1996, for a total of 323 observations. $\rho 1$,$\rho 2$ and w1, w2 are the correlation parameters and weights in copula 1 and 2, respectively. L represents the log likelihood function evaluated at the estimated parameters while AIC and BIC are as defined in the paper.

Table 3: Unconditional Model, Sample B

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US
$\rho 1$	0.708	0.205	0.356	0.520	-0.129	0.396	0.521	-0.893	-0.251	0.618
	(0.045)	(0.074)	(0.153)	(0.045)	(0.155)	(0.131)	(0.144)	(0.091)	(0.158)	(1.096)
$\rho 2$	0.883	0.867	0.780	0.905	0.573	0.770	0.674	0.441	0.598	0.636
	(0.027)	(0.055)	(0.031)	(0.053)	(0.057)	(0.035)	(0.082)	(0.048)	(0.057)	(0.650)
w1	0.433	0.756	0.193	0.812	0.286	0.245	0.358	0.076	0.312	0.374
w2	0.567	0.245	0.807	0.188	0.715	0.755	0.642	0.924	0.688	0.626
L	169.320	26.085	108.080	69.310	26.154	98.868	75.544	22.696	24.204	78.631
AIC	-332.640	-46.170	-210.160	-132.620	-46.308	-191.736	-145.088	-39.392	-42.408	-151.262
BIC	-321.225	-34.755	-198.745	-121.205	-34.893	-180.321	-133.673	-27.977	-30.993	-139.847

Table 3 presents parameter estimates from the two-component model. Standard errors are in parentheses. The frequency is weekly. The sample period is March 20, 1996 to May 29, 2002, for a total of 332 observations. $\rho 1$,$\rho 2$ and w1, w2 are the correlation parameters and weights in copula 1 and 2, respectively. L represents the log likelihood function evaluated at the estimated parameters, while AIC and BIC are as defined in the paper.

Table 4: Hidden Markov Model

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US
$\rho 1$	0.508	0.169	0.175	0.364	0.104	0.215	0.369	-0.076	-0.248	-0.169
	(0.238)	(0.100)	(0.152)	(0.097)	(0.122)	(0.143)	(0.120)	(0.109)	(0.091)	(0.175)
$\rho 2$	0.836	0.789	0.766	0.671	0.733	0.771	0.685	0.612	0.528	0.601
	(0.030)	(0.026)	(0.031)	(0.043)	(0.031)	(0.026)	(0.042)	(0.041)	(0.050)	(0.052)
π_{11}	0.712	0.688	0.870	0.738	0.777	0.802	0.975	0.206	0.255	0.237
	(0.047)	(0.026)	(0.035)	(0.027)	(0.024)	(0.034)	(0.009)	(0.028)	(0.033)	(0.063)
π_{12}	0.287	0.310	0.129	0.261	0.221	0.197	0.025	0.790	0.743	0.762
	(0.047)	(0.026)	(0.035)	(0.027)	(0.024)	(0.034)	(0.009)	(0.028)	(0.033)	(0.063)
π_{21}	0.099	0.784	0.026	0.236	0.369	0.071	0.020	0.449	0.335	0.064
	(0.019)	(0.032)	(0.008)	(0.024)	(0.035)	(0.013)	(0.008)	(0.029)	(0.026)	(0.012)
π_{22}	0.899	0.216	0.972	0.762	0.631	0.927	0.977	0.551	0.664	0.935
	(0.019)	(0.032)	(0.008)	(0.024)	(0.035)	(0.013)	(0.008)	(0.029)	(0.026)	(0.012)
L	311.910	95.581	235.450	120.050	93.553	213.530	127.120	95.343	77.947	130.930
AIC	-615.820	-183.162	-462.900	-232.100	-179.106	-419.060	-246.240	-182.686	-147.894	-253.860
BIC	-597.937	-165.279	-445.017	-214.217	-161.223	-401.177	-228.357	-164.803	-130.011	-235.977

Table 4 presents estimated parameters for a hidden markov model with two normal copulas. Standard errors are reported in parentheses. The frequency is weekly. The sample period is January 10, 1990 to May 29, 2002, for a total of 646 observations. $\rho 1$,$\rho 2$ are the correlation parameters in copula 1 and 2, and the parameters π_{11}, π_{12}, π_{21}, π_{22} are transition probabilities. L is the log likelihood evaluated at the estimated parameters.

Table 5: Hidden Markov Model, Normal-RG

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US
ρ	0.861	0.202	0.773	0.439	-0.152	0.791	0.457	0.666	0.717	0.633
	(0.027)	(0.144)	(0.032)	(0.076)	(0.197)	(0.023)	(0.088)	(0.042)	(0.038)	(0.064)
b	1.976	1.437	1.292	2.474	1.334	1.341	2.166	1.099	1.081	1.257
	(0.359)	(0.148)	(0.149)	(0.208)	(0.145)	(0.122)	(0.183)	(0.081)	(0.077)	(0.180)
π_{11}	0.706	0.453	0.968	0.959	0.913	0.896	0.940	0.248	0.425	0.695
	(0.035)	(0.031)	(0.009)	(0.009)	(0.042)	(0.017)	(0.011)	(0.028)	(0.040)	(0.025)
π_{12}	0.292	0.545	0.030	0.039	0.086	0.102	0.059	0.752	0.574	0.303
	(0.035)	(0.031)	(0.009)	(0.009)	(0.042)	(0.017)	(0.011)	(0.028)	(0.040)	(0.025)
π_{21}	0.144	0.431	0.095	0.125	0.005	0.149	0.164	0.582	0.237	0.726
	(0.019)	(0.028)	(0.025)	(0.028)	(0.003)	(0.024)	(0.029)	(0.029)	(0.023)	(0.035)
π_{22}	0.855	0.568	0.905	0.873	0.993	0.850	0.833	0.415	0.761	0.273
	(0.019)	(0.028)	(0.025)	(0.028)	(0.003)	(0.024)	(0.029)	(0.029)	(0.023)	(0.035)
L	302.690	54.142	231.270	134.760	54.892	209.410	129.950	86.026	73.436	125.680
AIC	-597.380	-100.284	-454.540	-261.520	-101.784	-410.820	-251.900	-164.052	-138.872	-243.360
BIC	-579.497	-82.401	-436.657	-243.637	-83.901	-392.937	-234.017	-146.169	-120.989	-225.477

Table 5 presents estimated parameters for a hidden markov model with normal and Rotated Gumbel components. The frequency is weekly. The sample period is January 10, 1990 to May 29, 2002, for a total of 646 observations. ρ and b are the dependence parameters in the normal and rotated gumbel copulas and the parameters π_{11}, π_{12}, π_{21} and π_{22} are transition probabilities. L is the log likelihood function evaluated at the estimated parameters.

Table 6: Kolmogorov-Smirnov and Anderson-Darling Statistics

Panel A: Kolmogorov-Smirnov

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US	Average
No(2)	0.36514	0.093576	0.23595	0.017811	0.07916	0.17759	0.024982	0.03235	0.016276	0.032291	
rank	3	3	3	3	3	3	3	2	2	2	2.7
NoRG(2)	0.15403	0.01913	0.19251	0.01241	0.01700	0.15585	0.01601	0.04064	0.05140	0.03328	
rank	2	2	2	2	1	2	2	3	3	3	2.2
Best Unc.	0.01902	0.01696	0.01663	0.01070	0.01700	0.01542	0.01587	0.01440	0.01286	0.01519	
rank	1	1	1	1	1	1	1	1	1	1	1.0

Table 6 presents estimates of the Kolmogorov-Smirnov and Anderson-Darling statistics, as described in the text. No(2) and NoRG(2) denote the two component models with two normal copulas, and one normal plus one Rotated Gumbel copula, respectively. Best Unc. Denotes the best unconditional model from a set of mixed and single copulas. Rank is the relative performance of the model within a specific country pair. The number 1 signifies the best performance, that is, the smallest statistic.

Panel B: Average Kolmogorov-Smirnov

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US	Average
No(2)	0.015641	0.004974	0.008055	0.003	0.006178	0.006668	0.005327	0.003817	0.004475	0.004232	
rank	3	3	3	3	3	2	3	3	3	3	2.9
NoRG(2)	0.00991	0.00399	0.00745	0.00286	0.00436	0.00700	0.00441	0.00370	0.00408	0.00363	
rank	2	2	2	1	2	3	2	2	2	2	2.0
Best Unc.	0.00456	0.00350	0.00343	0.00286	0.00421	0.00490	0.00440	0.00369	0.00370	0.00359	
rank	1	1	1	1	1	1	1	1	1	1	1.0

Panel C: Anderson-Darling

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US	Average
No(2)	3.83870	0.81805	1.16560	0.74597	0.79508	0.82077	0.72728	0.78887	0.55847	0.59283	
rank	3	3	3	3	3	3	3	3	3	1	2.8
NoRG(2)	1.59880	0.62199	0.95104	0.59039	0.45903	0.66106	0.50825	0.64064	0.49037	0.77410	
rank	2	2	2	2	1	2	2	2	2	3	2.0
Best Unc.	0.31842	0.50966	0.45454	0.58997	0.45903	0.51524	0.50378	0.50982	0.46294	0.61069	
rank	1	1	1	1	1	1	1	1	1	2	1.1

Panel D: Average Anderson-Darling

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US	Average
No(2)	0.084555	0.047941	0.044693	0.029937	0.059486	0.050191	0.041506	0.04412	0.044064	0.034433	
rank	3	3	3	3	3	3	3	3	3	3	3.0
NoRG(2)	0.04371	0.03344	0.03562	0.02474	0.03758	0.03755	0.03563	0.03589	0.03548	0.02878	
rank	2	2	2	1	1	2	2	2	2	2	1.8
Best Unc.	0.02553	0.03226	0.02520	0.02474	0.03758	0.03204	0.03562	0.03589	0.03325	0.02848	
rank	1	1	1	1	1	1	1	1	1	1	1.0

Table 7: Model Comparison

Panel A: AIC

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US	Average
No(1)	-537.800	-79.986	-371.820	-200.640	-75.174	-305.960	-210.860	-87.646	-55.394	-213.880	
rank	5	5	5	3	5	5	3	5	5	5	4.6
No(2)	-547.880	-85.284	-383.500	-196.720	-83.228	-324.240	-207.100	-97.142	-64.808	-215.960	
rank	4	4	4	5	4	4	5	4	4	4	4.2
HMM-No	-615.820	-183.162	-462.900	-232.100	-179.106	-419.060	-246.240	-182.686	-147.894	-253.860	
rank	1	1	1	2	1	1	2	1	1	1	1.2
HMM-NoRG	-597.380	-100.284	-454.540	-261.520	-101.784	-410.820	-251.900	-164.052	-138.872	-243.360	
rank	2	2	2	1	2	2	1	2	2	2	1.8
Best Unc.	-565.120	-85.706	-386.720	-200.640	-89.242	-333.380	-211.240	-99.344	-68.646	-216.840	
rank	3	3	3	3	3	3	3	3	3	3	3.0

Table 7 compares model performance within each country pair, using AIC and BIC. No(1) and No(2) denote results from the single and two-component normal copula models, respectively. HMM-No and HMM-NoRG denote results from the hidden markov models with two normal copulas and with a normal and rotated gumbel copula, respectively. Best Unc. denotes the best unconditional model from a set of mixed and single copulas. Rank is the performance of each model within the specific country pair. The Number 1 denotes the best performance, that is, the lowest AIC or BIC.

Panel B: BIC

	FR-DE	FR-JP	FR-UK	FR-US	DE-JP	DE-UK	DE-US	JP-UK	JP-US	UK-US	Average
No(1)	-533.329	-75.515	-367.349	-196.169	-70.703	-301.489	-206.389	-83.175	-50.923	-209.409	
rank	5	4	5	3	4	5	4	5	5	3	4.3
No(2)	-534.468	-71.872	-370.088	-183.308	-69.816	-310.828	-193.688	-83.730	-51.396	-202.548	
rank	4	5	4	5	5	4	5	4	4	4	4.4
HMM-No	-597.937	-165.279	-445.017	-214.217	-161.223	-401.177	-228.357	-164.803	-130.011	-235.977	
rank	1	1	1	2	1	1	2	1	1	1	1.2
HMM-NoRG	-579.497	-82.401	-436.657	-243.637	-83.901	-392.937	-234.017	-146.169	-120.989	-225.477	
rank	2	2	2	1	3	2	1	2	2	2	1.9
Best Unc.	-560.649	-81.235	-375.218	-196.169	-84.771	-322.409	-206.389	-92.369	-59.704	-209.409	
rank	3	3	3	3	2	3	3	3	3	3	2.9